

QUE PRONÚNCIA DAR À LEITURA?

Maria Carlota ROSA
 Carlos Alexandre GONÇALVES
 Ruth MONSERRAT
 (UFRJ/Faculdade de Letras)

RESUMO: Parte de uma pesquisa que visa a desenvolver um sistema de síntese text-to-speech para o português do Brasil, este trabalho baseia-se em Ramos (1997) e em Soares (1979), sugerindo uma pronúncia neutra para o sintetizador.

PALAVRAS-CHAVE: Síntese da fala; sotaque; variação fonológica.

0. Introdução

Estabelecer como soa cada unidade do sistema de escrita do português — sejam essas unidades grafemáticas ou extragrafemáticas — pressupõe pelo menos duas decisões prévias: (a) que sotaque será atribuído à leitura; e (b) que características próprias da leitura deverão ser incluídas. O primeiro aspecto leva em conta não apenas a escolha de uma variedade regional, mas também o grau de formalidade; o segundo, padrões não necessariamente presentes na fala (REBOLLO COUTO, 1999).

Para um *software* que, em princípio, poderá vir a ser utilizado em qualquer parte do Brasil, põe-se o desiderato de uma pronúncia neutra. Para determinar essa *pronúncia neutra*, retomamos os trabalhos de RAMOS (1997) e SOARES (1977), além de consultarmos informalmente locutores profissionais. A opção por um registro formal (ao menos para um produto inicial) impediu, por exemplo, a inclusão de fenômenos como a monotongação de ditongos decrescentes diante de consoante palatal na sílaba seguinte, acompanhado ou não do ensurdecimento ou supressão da vogal átona final, como em [ba|o] ou [pe|e], possíveis realizações para <baixo> e <peixe>, nesta ordem.

1. A variante escolhida

Uma das variantes do português do Brasil é a que se convencionou chamar *português padrão* (cf. Cunha, 1970). É uma variedade de língua que passou por um processo de codificação, refletido em dicionários de grande porte, no *Vocabulário Ortográfico* da Academia Brasileira de Letras, em inúmeras obras gramaticais e em vasta literatura, com variação relativamente pequena no que toca à gramática e ao vocabulário. Pelo tipo de proposta do Projeto, o português padrão — a variedade de língua empregada na imprensa, nos telejornais, ensinada na escola e falada pelos indivíduos instruídos (Trudgill, 1983) — foi tomado por base.

O português padrão engloba tanto variantes coloquiais como variantes formais. A opção pelas variantes formais do português padrão, ao menos para a etapa inicial do Projeto, justifica-se em função do *corpus* testado nessa fase. No atual estágio da pesquisa, está sendo utilizada uma base de dados eletrônica, o CETENFolha, *corpus* de cerca de 24 milhões de palavras do português do Brasil, construído com base em extratos de notícias de todas as edições do ano de 1994 do jornal *Folha de São Paulo*. Os *e-mails*, por exemplo, ficam para uma fase posterior da pesquisa.

Se houve consenso imediato quanto à escolha da variante, o mesmo aconteceu em relação à pronúncia que o português padrão receberia. A literatura lingüística de três ou mais décadas atrás apontava a pronúncia do Rio de Janeiro como a ideal, aquela que deveria ser ensinada na escola, como afirmava Cunha (1970:83):

No Brasil, onde, por decisão do Conselho Federal de Educação, a pronúncia modelo (a que deve ser ensinada nas escolas) é basicamente a do Rio de Janeiro — já assim reconhecida por dois congressos¹³¹ — a preocupação será inversa [**à preocupação em Portugal, onde as vogais átonas estão sendo obscurecidas –mcr**], a de proteger as consoantes, evitando, por exemplo, a evolução das africadas [tʃ] e [dʒ] (tipo: [noy tʃi] ‘noite’, [dʒu] ‘ódio’).

Tal quadro não se mantém nos dias de hoje. A pronúncia carioca é vista por brasileiros de outras regiões como “chiada”, mas especialmente como muito nasalada, mesmo em contextos em que não haja consoantes ou vogais nasais. Pesquisa realizada por Ramos (1997), que retoma Almeida (1979), confirma para quatro, das cinco capitais focalizadas (Belo Horizonte, Florianópolis, Porto Alegre e João Pessoa, mas não o Rio de Janeiro), a progressiva perda de prestígio dessa pronúncia.

De acordo com resultados de Ramos (1997), em cerca de 20 anos, a aceitação do dialeto carioca como modelo para os demais falantes do País baixara de 19,9% para 8,5%. Ao mesmo tempo, Ramos aponta a pronúncia empregada no *Jornal Nacional*, o telejornal com os maiores índices de audiência no País, como aquela que falantes de diferentes regiões brasileiras identificam, em maior ou menor grau, com a sua própria: algo que, por conseguinte, pode ser denominado “pronúncia neutra”. Como contraponto, o dialeto carioca — nos resultados parciais obtidos pela pesquisa de Ramos em BH — foi considerado o mais marcado e, portanto, o mais distante do padrão da mídia.

Os resultados em Ramos (1997) mostram que a perda de prestígio não é percebida pelos cariocas. Não obstante a rejeição apontada pela autora nas demais capitais, 45,9% dos entrevistados cariocas (o maior percentual dentre os entrevistados) afirmaram que o *Jornal Nacional* refletia a fala de sua região, ao contrário, por exemplo, dos paraibanos: apenas 8% deles consideraram a fala de sua região refletida nesse noticiário. Mas o que a pronúncia do *Jornal Nacional* teria de neutra? Afinal, conta com uma carioca, a âncora Fátima Bernardes, entre seus apresentadores diários.

2. Uma pronúncia neutra para o português padrão

Entrevistamos informalmente¹³² um âncora da mesma emissora, Sidney Resende, um dos responsáveis por um telejornal local, o *RJTV*. Perguntamos a ele se havia diferenças entre assumir um noticiário local, como o *RJTV*, e um para todo o território, como o *Jornal Nacional*.

Acredito que exista diferença, sim. Num jornal local a forma da sua gente se expressar poderia ser adotada sem críticas. Já num jornal nacional uma gíria, por exemplo, poderia não ser entendida. Um sotaque muito forte pode, quem sabe, até prejudicar a compreensão de uma palavra ou frase.

A resposta de Resende vai ao encontro do comentário de Cunha, anteriormente citado, quando exemplifica o que seria por demais carioca, os “erres” e os “esses”:

Procuo dar uma pronúncia “neutra”. Se é que isso é possível. Ou seja, evito carregar nos erres, que é muito comum no Rio. Por que faço isso?? É para dar aos brasileiros uma narrativa que mereça a aceitação geral.

¹³¹ *Anais do Primeiro Congresso da Língua Nacional Cantada*, São Paulo: 1938; *Anais do Primeiro Congresso Brasileiro de Língua Falada no Teatro*, Rio de Janeiro: 1958.

¹³² Por *informalmente*, entenda-se ‘por e-mail’, datados todos de 06.07.2002.

A resposta de Resende vai também ao encontro do que fora levantado por Soares (1977) como a norma ideal entre locutores de noticiários que liam seus textos, cuja profissão era exercida no Rio de Janeiro.

3. Questões fonético-fonológicas analisadas

Dentre as variações fonológicas do português brasileiro, destacam-se, por seu caráter dialetal, (a) a articulação das vogais pretônicas; (b) a manifestação fonética de consoantes em posição de coda (em especial, de “erres” e “esses”); e (c) a nasalização de vogais orais diante de consoantes nasais. A realização desses segmentos já foi adotada como critério para a delimitação de áreas regionais e para a diferenciação de falares e subfalares (cf., p. ex., Nascentes, 1953, e Houaiss, 1958). Além disso, serviu de base para a caracterização do “português padrão” (cf. Cunha, 1970) e para o estabelecimento de uma “pronúncia radiofônica neutra” (cf. Soares, 1977). Por esses motivos, um sintetizador que possa ser utilizado em qualquer parte do Brasil deve ser capaz de neutralizar essas diferenças, de modo a oferecer ao usuário uma pronúncia que, de certo modo, se identifique com a sua.

Os casos de variação fonológica acima mencionados dependem das noções de sílaba e acento como um *a priori* para sua aplicação, uma vez que fazem referência à posição de um segmento no nível da sílaba (se em coda ou em *onset*) ou às relações de tonicidade. Dessa forma, para que o *software* converta texto em fala de modo inteligível e natural, é necessário reconhecer, a partir de uma seqüência de grafemas, o que é uma sílaba, como ela se constitui internamente e qual é a sua posição quanto ao acento da palavra.

Para a identificação da sílaba¹³³ tônica a partir do sistema ortográfico do português, propusemos um algoritmo em forma de árvore de decisões (Odell, 1995). Em linhas gerais, tal recurso consiste numa lista hierárquica de perguntas com respostas simples, sim/não, formuladas a partir do estabelecimento de regras categóricas. As regras são propostas de modo a determinar que tipo de comportamento articulatório deve-se associar às palavras e às letras do texto escrito. O nó inicial, aquele que representa qualquer contexto, é a noção de *palavra*, definida graficamente, isto é, como elemento entre espaços e/ou pontuação. A seleção dos nós subsequentes dependerá da resposta, *sim* ou *não*, sobre o contexto, até que não mais caibam perguntas (v. Figura 1).

¹³³ Para a identificação das sílabas o ponto de partida é um programa modificado por Fernando Gil Resende Jr., Felipe Barbosa e Guilherme Pinto.

Com o propósito de privilegiar pronúncias menos marcadas, optamos por neutralizar essas diferenças, associando os grafemas <e> e <o> sempre aos fones [e] e [o], nesta ordem. Basicamente, quatro foram os motivos que nos levaram a essa escolha: (a) na leitura, a articulação *default* dessas vogais é a de média fechada (cf. Silva, 1991); (b) pronúncias mais abertas, ao que tudo indica, ainda parecem estar vinculadas aos subfalares da região Nordeste (cf. Houaiss, 1958); (c) os locutores entrevistados por Soares (1977) recomendam o uso das médias de 2º. grau, advertindo, ainda, que “*não se deve pronunciar [u] onde se escreve o (como tomate, botei, porque), na posição pretônica*” (cf. Soares, 1977: 6); e, por fim, (d) nem sempre é previsível o alteamento das pretônicas por harmonização vocálica, uma vez que tônicas com [i, u] podem não exercer qualquer influência sobre a átona que a precede (cf. *pecúlio*, nunca realizado como *picúlio*, e *coluna*, dificilmente pronunciado *culuna*), fato que inviabilizaria o estabelecimento de uma regra categórica para a árvore de decisões.

Em relação às médias pretônicas, ainda constituem resíduos para o sintetizador palavras como *belíssima*, *novamente* e *velhinha*. Em todas elas, a pretônica nunca se realiza como fechada, contrariando, portanto, a tendência geral à associação dos grafemas <e, o> aos fones [e, o], respectivamente. A abertura dessas átonas se justifica por questões de natureza morfológica e, por isso, a terminação da palavra constitui informação relevante para a eventual existência de vogais abertas em palavra derivadas. Como *-mente*, *-íssimo* e *-inho* são os únicos sufixos que não promovem o fechamento da pretônica, palavras derivadas podem apresentar tanto vogais abertas (*certamente*, *períssimo*, *novinha*) quanto fechadas (*tolamente*, *cedíssimo*, *lobinho*), o que constitui problema, levando-se em conta que o *input* é um texto escrito e não há pistas ortográficas que indiquem o timbre da pretônica. Estamos viabilizando a criação de um processador morfológico para resolver problemas desse tipo.

Outros sons sujeitos à controvérsia quanto às normas de pronúncia são os “erres” e os “esses” em posição implosiva (cf. Soares, 1977). Para garantir realizações mais naturais, optamos por registrar a assimilação de sonoridade característica desses dois segmentos não-especificados quanto ao vozeamento. Sem sonoridade inerente, tais elementos são influenciados pelo ambiente fônico e se articulam vozeados diante de consoantes sonoras (cf. *vesgo* e *carga*) e desvozeados diante de surdas (cf. *casca* e *marca*). Como essa regra de assimilação não se restringe a uma área dialetal específica, aplicando-se em quase todo o território brasileiro, consideramos oportuno registrar essa diferença fonética.

Se, por um lado, as variações de sonoridade devem ser assinaladas, com vistas a veicular uma pronúncia mais natural, por outro, as de ponto de articulação têm de ser neutralizadas, para que se chegue a uma pronúncia mais neutra. Na *seção 1*, mostramos que o “chiado” constitui traço característico do português do Rio de Janeiro e que os locutores tendem a evitar a articulação alveopalatal ([s, z]) por perceberem nela forte associação com o falar carioca. Dessa forma, optamos pelas realizações alveolares ([s, z]), consideradas neutras e, portanto, menos marcadas do ponto-de-vista diatópico. Para os locutores, há, ainda, uma razão técnica para preferir o uso das alveopalatais: sons chiados “*soam mal ao microfone*” e, por isso mesmo, “*devem ser descartados*” (Soares, 1977: 4).

De todos os sons do português, o que apresenta realizações mais variadas é, sem dúvida, o “erre” implosivo. Por exemplo, esse segmento pode-se realizar como fricativa velar (p. ex., Rio de Janeiro), vibrante múltipla ápico-alveolar (p. ex., Rio Grande do Sul), retroflexa (interior de São Paulo e Minas Gerais), fricativa glotal (p. ex., Belo Horizonte), tepe alveolar (p. ex. São Paulo capital) ou mesmo como um zero fonético em final de palavra (como *cantá*, por *cantar*). Determinadas realizações são mais marcadas que outras e, por isso mesmo, mais diretamente vinculadas a um tipo específico de falar. Ao que tudo indica, é bastante estigmatizada a pronúncia retroflexa e a carioca, por sua vez, sentida como muito “*intensa e por demais gutural*” (Christófaros Silva, 1999: 23).

Por ser menos audível, do ponto-de-vista perceptivo, e mais débil, do ponto-de-vista articulatório (Auler, 1993: 25), a realização aspirada do “erre” (como fricativa glotal) foi adotada na pesquisa, por ser considerada a mais neutra e a menos vinculada a um determinado dialeto. A realização do “erre” como zero fonético foi descartada não só por ocorrer apenas em final de palavra, mas, principalmente, porque: (a) optamos por privilegiar as variedades mais formais do português padrão; e (b) estamos trabalhando

basicamente com leitura. A esse respeito, recomendam as normas de pronúncia radiofônica que “*não se deve omitir sons no meio ou no fim das palavras, especialmente o r final*” (Soares, 1977: 7).

Por fim, cabem alguns comentários sobre o que se conhece na literatura por nasalidade fonética. Antes de consoantes nasais, vogais orais tendem a assimilar o traço de nasalidade do elemento seguinte, sendo articuladas, pois, como nasais. Em determinados falares, como o baiano, essa é a única condição que engatilha o processo e, portanto, numa palavra como *banana*, as duas primeiras vogais são igualmente nasalizadas. Em outros falares, diferentemente, o acento constitui informação adicional à regra de nasalização. Por exemplo, no Rio de Janeiro só se nasalizam as vogais que, precedendo uma consoante nasal na sílaba seguinte, aparecem na posição tônica. Assim, no Rio de Janeiro, a mesma palavra só apresenta uma vogal nasalizada (a acentuada), recebendo a primeira uma articulação nitidamente oral.

Com o objetivo de atribuir uma pronúncia natural e neutra à leitura de qualquer texto escrito em português, também neutralizamos as diferenças de nasalização, formulando uma regra que converte em nasais apenas as vogais tônicas. Acreditamos, com isso, que estamos contemplando, mais uma vez, as pronúncias menos marcadas, uma vez que a nasalidade de átonas pretônicas parece ser traço característico dos falares e subfalares do Nordeste, ainda recorrendo à divisão estabelecida por Nascentes (1953). Apesar de dar conta de muitos casos, a regra não contempla formas como *c[ã]ninha* e *hum[ã]namente*. Nesses dois exemplos, a nasalidade independe do acento e, ao que tudo indica, é categórica em todos os falares do português brasileiro. Novamente, deparamo-nos com a necessidade de implementar um processador morfológico, pois há sufixos que são sensíveis à nasalidade da palavra derivante, como *-inho* e *-mente* (cf. *c[ã]minha*, diminutivo de *cama*, mas *c[a]minha*, do verbo *caminhar*), ao lado de sufixos que ignoram esse tipo de informação fonética (cf., p. ex., *c[ã]no*, mas *c[a]nal*, *c[a]nudo*, *c[a]neta*).

Referências bibliográficas:

- ALMEIDA, M. J. A. 1979. *Étude sur les attitudes linguistiques au Brésil*. Université de Montreal. 1979.
- AULER, Mônica. 1993. *A aspiração do “esse” final: uma regra variável*. Dissertação de Mestrado em Linguística. Rio de Janeiro: UFRJ/Faculdade de Letras, mimeo.
- CETENFolha (Corpus de Extractos de Textos Electrónicos NILC/Folha de S. Paulo). <http://acdc.linguateca.pt/cetenfolha/>.
- CHRISTÓFARO SILVA, Thaís. 1999. *Fonética e Fonologia do português*. São Paulo: Contexto.
- CUNHA, Celso. 1970. *Língua portuguesa e realidade nacional*. 2ª. ed. atualiz. Rio de Janeiro: Tempo Brasileiro.
- HOUAISS, Antonio. 1958. Tentativa de descrição do sistema vocálico do português culto na área dita carioca. *Anais do Congresso Brasileiro de língua falada no teatro I* (Salvador, 1956). Rio de Janeiro: MEC.
- NASCENTES, Antenor. 1953. *O linguajar carioca*. Rio de Janeiro: Simões.
- ODELL, Julian James. 1995. *The use of context in large vocabulary speech recognition*. Cambridge, Gr. Brit.: University of Cambridge/Queen’s College. mimeo.
- RAMOS, Jânia. 1997. Avaliação de dialetos brasileiros: o sotaque. *Revista de Estudos da Linguagem*. Belo Horizonte: UFMG, 6(5)103-125.
- REBOLLO COUTO, Letícia. 1999. *Rythme et mélodie de la parole en Espagnol et en Portugais du Brésil: analyse acoustique et journaux télévisés*. Strasbourg: Université Marc Bloch. mimeo
- SILVA, Ademar. 1991. *Redação: a escrita espontânea*. São Paulo: Ática.
- SOARES, Maria Aparecida Botelho Pereira. 1977. *Normas de pronúncia radiofônica no Rio de Janeiro*. UFRJ: Museu Nacional.
- TRUDGILL, P. 1983. *Sociolinguistics: An introduction to language and society*. Harmondsworth: Penguin Books.